

3/4



# 3

ENTERED OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/931,457A

DATE: 03/06/2002 P. 6

TIME: 10:18:32

Input Set : A:\BB1116 US CIP Seq. list correct.txt

Output Set: N:\CRF3\03062002\I931457A.raw

```

4 <110> APPLICANT: Falco, S. Carl
5   Allen, Stephen M.
7 <120> TITLE OF INVENTION: Plant Amino Acid Biosynthetic Enzymes
9 <130> FILE REFERENCE: BB1116 US CIP
11 <140> CURRENT APPLICATION NUMBER: 09/931,457A
C--> 12 <141> CURRENT FILING DATE: 2002-02-22
14 <150> PRIOR APPLICATION NUMBER: 09/424,976
15 <151> PRIOR FILING DATE: 1999-12-02
17 <150> PRIOR APPLICATION NUMBER: 60/065,385
18 <151> PRIOR FILING DATE: 1997-11-12
20 <150> PRIOR APPLICATION NUMBER: 60/049,406
21 <151> PRIOR FILING DATE: 1997-06-12
23 <160> NUMBER OF SEQ ID NOS: 72
25 <170> SOFTWARE: Microsoft Office 97
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 826
29 <212> TYPE: DNA
30 <213> ORGANISM: Oryza sativa
32 <400> SEQUENCE: 1
33 tgggtaccgcc acgccaaggt ggtaaggatg gttgtcagca cttaccaagc agcaagtggg      60
34 gctggggctg cggccatgga agaactcaaa cttcaaaactc aagaggtctt ggcggggaaa      120
35 gcaccaacat gcaacatttt cagtcagcag tatgctttta atatattttc acataatgca      180
36 ccaattgttg aaaatgggta caatgaggag gagatgaaga tggatgaaga gaccagaaaa      240
37 atctggaatg ataaagatgt gaaggtaact gcaacctgca tacgagttcc tgtgatgcgt      300
38 gcacatgctg aaagtgtgaa tctacagttt gaaaagccac ttgatgagga tactgcaagg      360
39 gaaatcttga gggcagctga aggtgttacc attattgatg accgtgcttc caatcgcttc      420
40 cccacacctc ttgaggtatc ggataaagat gatgtagcag tgggtagaat tcgtcaggat      480
41 ttgtcgcaag atgataacaa agggctggac atatttgttt gtggagatca aatacgtaaa      540
42 ggtgctgcac tcaatgctgt gcagattgct gaaatgctac tcaagtgatt ttcttttctg      600
43 tacctttctc tccttgcccc tctttgctct agtcattggt tgacggatgt actctgggta      660
44 gtatgagatc aattttgatc atcttttgta atctatatc ctagtgaaat aaatgtaaaa      720
45 cggttttgct ctatcttctg cacaagtgtg gaagaaatct gaaattggga aattggagtg      780
46 tggcccttgt tcaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa      826
48 <210> SEQ ID NO: 2
49 <211> LENGTH: 195
50 <212> TYPE: PRT
51 <213> ORGANISM: Oryza sativa
53 <400> SEQUENCE: 2
54 Trp Tyr Arg His Ala Lys Val Val Arg Met Val Val Ser Thr Tyr Gln
55 1           5           10           15
57 Ala Ala Ser Gly Ala Gly Ala Ala Ala Met Glu Glu Leu Lys Leu Gln
58           20           25           30
60 Thr Gln Glu Val Leu Ala Gly Lys Ala Pro Thr Cys Asn Ile Phe Ser

```

## RAW SEQUENCE LISTING

DATE: 03/06/2002

PATENT APPLICATION: US/09/931,457A

TIME: 10:18:32

Input Set : A:\BB1116 US CIP Seq. list correct.txt

Output Set: N:\CRF3\03062002\I931457A.raw

```

61          35          40          45
63 Gln Gln Tyr Ala Phe Asn Ile Phe Ser His Asn Ala Pro Ile Val Glu
64          50          55          60
66 Asn Gly Tyr Asn Glu Glu Glu Met Lys Met Val Lys Glu Thr Arg Lys
67 65          70          75          80
69 Ile Trp Asn Asp Lys Asp Val Lys Val Thr Ala Thr Cys Ile Arg Val
70          85          90          95
72 Pro Val Met Arg Ala His Ala Glu Ser Val Asn Leu Gln Phe Glu Lys
73          100          105          110
75 Pro Leu Asp Glu Asp Thr Ala Arg Glu Ile Leu Arg Ala Ala Glu Gly
76          115          120          125
78 Val Thr Ile Ile Asp Asp Arg Ala Ser Asn Arg Phe Pro Thr Pro Leu
79          130          135          140
81 Glu Val Ser Asp Lys Asp Asp Val Ala Val Gly Arg Ile Arg Gln Asp
82 145          150          155          160
84 Leu Ser Gln Asp Asp Asn Lys Gly Leu Asp Ile Phe Val Cys Gly Asp
85          165          170          175
87 Gln Ile Arg Lys Gly Ala Ala Leu Asn Ala Val Gln Ile Ala Glu Met
88          180          185          190
90 Leu Leu Lys
91          195
93 <210> SEQ ID NO: 3
94 <211> LENGTH: 875
95 <212> TYPE: DNA
96 <213> ORGANISM: Triticum aestivum
98 <400> SEQUENCE: 3
99 cctcatgggt gtcacgccgc tgcacgccca cgccaagggtg aaaaggatgg ttgtcagcac      60
100 ataccaagca gcaagtgggtg ctgggtgctgc agccatggaa gaactcaaac ttcagactcg      120
101 agaggtcttg gaaggaaagc caccaacctg taacattttc agtcaacagt atgcttttaa      180
102 tatattttcg cataatgcac ctattgttga aaatggctat aatgaggaag agatgaaaat      240
103 ggtgaaggag accagaaaaa tctggaatga caaggatgta agagtaactg caacttgtat      300
104 acgggttccct acgatgcgcg cgcatgccga aagcgtgaat ctacagtttg aaaagccact      360
105 tgatgaggac actgccagag aaatcttgag ggcagctcct ggtgttacca ttagtgacga      420
106 ccgtgctgcc aaccgcttcc ctacaccact ggaggtatcg gataaagatg acgtatcagt      480
107 tggtaggatt cgccaggact tgtcacaaga tgataacaga gggttggagt tatttgtctg      540
108 tggagaccag atacgtaaag gcgcgcgcgt gaacgctgtg cagattgctg aaatgctact      600
109 gaagtgaccg cctttttacc attgtctcat gtgccacgtt gctctatcca ttgatggatt      660
110 gatgtactct agtcactttc aaccagttt tggtcgtcgt cttttttgta atctgtcaac      720
111 ctagcagaag aagtgtaaga cgggcttttag tcatctgttg cacacaaaag tgcagccaca      780
112 agtttagaaa aggaggggtt tcaactgttc ggattttgcc ttaggttgga ctttggtgca      840
113 agttgtcgtt tgtttcttga aagctggtct gctgt
115 <210> SEQ ID NO: 4
116 <211> LENGTH: 201
117 <212> TYPE: PRT
118 <213> ORGANISM: Triticum aestivum
120 <400> SEQUENCE: 4
121 Leu Met Ala Val Thr Pro Leu His Arg His Ala Lys Val Lys Arg Met
122 1          5          10          15
124 Val Val Ser Thr Tyr Gln Ala Ala Ser Gly Ala Gly Ala Ala Ala Met

```

## RAW SEQUENCE LISTING

DATE: 03/06/2002

PATENT APPLICATION: US/09/931,457A

TIME: 10:18:32

Input Set : A:\BB1116 US CIP Seq. list correct.txt

Output Set: N:\CRF3\03062002\I931457A.raw

```

125          20          25          30
127 Glu Glu Leu Lys Leu Gln Thr Arg Glu Val Leu Glu Gly Lys Pro Pro
128          35          40          45
130 Thr Cys Asn Ile Phe Ser Gln Gln Tyr Ala Phe Asn Ile Phe Ser His
131          50          55          60
133 Asn Ala Pro Ile Val Glu Asn Gly Tyr Asn Glu Glu Glu Met Lys Met
134 65          70          75          80
136 Val Lys Glu Thr Arg Lys Ile Trp Asn Asp Lys Asp Val Arg Val Thr
137          85          90          95
139 Ala Thr Cys Ile Arg Val Pro Thr Met Arg Ala His Ala Glu Ser Val
140          100          105          110
142 Asn Leu Gln Phe Glu Lys Pro Leu Asp Glu Asp Thr Ala Arg Glu Ile
143          115          120          125
145 Leu Arg Ala Ala Pro Gly Val Thr Ile Ser Asp Asp Arg Ala Ala Asn
146          130          135          140
148 Arg Phe Pro Thr Pro Leu Glu Val Ser Asp Lys Asp Asp Val Ser Val
149 145          150          155          160
151 Gly Arg Ile Arg Gln Asp Leu Ser Gln Asp Asp Asn Arg Gly Leu Glu
152          165          170          175
154 Leu Phe Val Cys Gly Asp Gln Ile Arg Lys Gly Ala Ala Leu Asn Ala
155          180          185          190
157 Val Gln Ile Ala Glu Met Leu Leu Lys
158          195          200

```

```

160 <210> SEQ ID NO: 5
161 <211> LENGTH: 457
162 <212> TYPE: DNA
163 <213> ORGANISM: Glycine max
165 <220> FEATURE:
166 <221> NAME/KEY: unsure
167 <222> LOCATION: (211)
168 <223> OTHER INFORMATION: n = A, C, G or T
170 <220> FEATURE:
171 <221> NAME/KEY: unsure
172 <222> LOCATION: (320)
173 <223> OTHER INFORMATION: n = A, C, G or T
175 <220> FEATURE:
176 <221> NAME/KEY: unsure
177 <222> LOCATION: (377)
178 <223> OTHER INFORMATION: n = A, C, G or T
180 <220> FEATURE:
181 <221> NAME/KEY: unsure
182 <222> LOCATION: (391) (392)
183 <223> OTHER INFORMATION: n = A, C, G or T
185 <220> FEATURE:
186 <221> NAME/KEY: unsure
187 <222> LOCATION: (410)
188 <223> OTHER INFORMATION: n = A, C, G or T
190 <220> FEATURE:
191 <221> NAME/KEY: unsure

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/931,457A

DATE: 03/06/2002

TIME: 10:18:32

Input Set : A:\BB1116 US CIP Seq. list correct.txt  
 Output Set: N:\CRF3\03062002\I931457A.raw

```

192 <222> LOCATION: (420)
193 <223> OTHER INFORMATION: n = A, C, G or T
195 <220> FEATURE:
196 <221> NAME/KEY: unsure
197 <222> LOCATION: (428)
198 <223> OTHER INFORMATION: n = A, C, G or T
200 <220> FEATURE:
201 <221> NAME/KEY: unsure
202 <222> LOCATION: (432)
203 <223> OTHER INFORMATION: n = A, C, G or T
205 <220> FEATURE:
206 <221> NAME/KEY: unsure
207 <222> LOCATION: (434)
208 <223> OTHER INFORMATION: n = A, C, G or T
210 <220> FEATURE:
211 <221> NAME/KEY: unsure
212 <222> LOCATION: (438)
213 <223> OTHER INFORMATION: n = A, C, G or T
215 <220> FEATURE:
216 <221> NAME/KEY: unsure
217 <222> LOCATION: (442)
218 <223> OTHER INFORMATION: n = A, C, G or T
220 <220> FEATURE:
221 <221> NAME/KEY: unsure
222 <222> LOCATION: (451)
223 <223> OTHER INFORMATION: n = A, C, G or T
225 <400> SEQUENCE: 5
226 gtctgtttta aaatccaaca cttaattctct ctcttcgcag cctaaaaatcc caatggcttc      60
227 actctctgtt ttgcgccaca accacctctt ctggggcccc ctcccgcccc gccccaagcc      120
228 cacctctccc tctctctcca ggatccgaat gtccctccgc gagaacggcc cctccatcgc      180
W--> 229 cgctcgtgggc gtcaccggcg ccgtcggcca ngagtctctc tccgtctctc ccgaccgcga      240
230 cttccctac cgctccattc atatgctggc ttccaagcgc tccgctggac gccgcatcac      300
W--> 231 cttcgaggac agggactacn tcttcaggag ctacgcgcgg agagtctcgc ggtgtcgaca      360
W--> 232 tcgcgtctct cagcgcnggg ggtccatcaa nnaagcattc ggaccatcgn cgtaaatacgn      420
W--> 233 gggacggncg tngncaanat anctccggtt nccttttg      457
235 <210> SEQ ID NO: 6
236 <211> LENGTH: 86
237 <212> TYPE: PRT
238 <213> ORGANISM: Glycine max
240 <400> SEQUENCE: 6
241 Met Ala Ser Leu Ser Val Leu Arg His Asn His Leu Phe Ser Gly Pro
242 1 5 10 15
244 Leu Pro Ala Arg Pro Lys Pro Thr Ser Ser Ser Ser Arg Ile Arg
245 20 25 30
247 Met Ser Leu Arg Glu Asn Gly Pro Ser Ile Ala Val Val Gly Val Thr
248 35 40 45
250 Gly Ala Val Gly Gln Glu Phe Leu Ser Val Leu Ser Asp Arg Asp Phe
251 50 55 60
253 Pro Tyr Arg Ser Ile His Met Leu Ala Ser Lys Arg Ser Ala Gly Arg

```

## RAW SEQUENCE LISTING

DATE: 03/06/2002

PATENT APPLICATION: US/09/931,457A

TIME: 10:18:32

Input Set : A:\BB1116 US CIP Seq. list correct.txt

Output Set: N:\CRF3\03062002\I931457A.raw

```

254 65          70          75          80
256 Arg Ile Thr Phe Glu Asp
257          85
259 <210> SEQ ID NO: 7
260 <211> LENGTH: 160
261 <212> TYPE: PRT
262 <213> ORGANISM: Legionella pneumophila
264 <400> SEQUENCE: 7
265 Met Ser Arg His Leu Asn Val Ala Ile Val Gly Ala Thr Gly Ala Val
266 1          5          10          15
268 Gly Glu Thr Phe Leu Thr Val Leu Glu Glu Arg Asn Phe Pro Ile Lys
269          20          25          30
271 Ser Leu Tyr Pro Leu Ala Ser Ser Arg Ser Val Gly Lys Thr Val Thr
272          35          40          45
274 Phe Arg Asp Gln Glu Leu Asp Val Leu Asp Leu Ala Glu Phe Asp Phe
275          50          55          60
277 Ser Lys Val Asp Leu Ala Leu Phe Ser Ala Gly Gly Ala Val Ser Lys
278 65          70          75          80
280 Glu Tyr Ala Pro Lys Ala Val Ala Ala Gly Cys Val Val Val Asp Asn
281          85          90          95
283 Thr Ser Cys Phe Arg Tyr Glu Asp Asp Ile Pro Leu Val Val Pro Gly
284          100         105         110
286 Ser Glu Ser Ser Ser Asn Arg Asp Tyr Thr Lys Arg Gly Ile Ile Ala
287          115         120         125
289 Asn Pro Asn Cys Ser Thr Ile Gln Met Val Val Ala Leu Lys Pro Ile
290          130         135         140
292 Tyr Asp Ala Val Gly Ile Ser Arg Ile Asn Val Ala Thr Tyr Gln Ser
293 145         150         155         160
296 <210> SEQ ID NO: 8
297 <211> LENGTH: 1054
298 <212> TYPE: DNA
299 <213> ORGANISM: Zea mays
301 <400> SEQUENCE: 8
302 atttaacgga aatgggaaga cactcgaaca tcttaaatta gctgctgaga gtggagtatt      60
303 tgtaaatgtg gatagcgaat ttgatttgga gaattattgtc agagctgcaa gagctactgg      120
304 aaagaaagtg cctgttttgc ttcgaataaa tccagatgtg gatccgcagg tacatcctta      180
305 tgttgccacg ggaataaaaa cgtctaaatt tgggatccgc aatgagaaat tgcaatgggt      240
306 tttggactct atcaagtcac acccgaaatga aatcaaactc gttggtgttc attgccatct      300
307 gggatctact attacaaagg ttgatattat cagagatgct gcagttctta tgctgaatta      360
308 tgtcgatgaa attcgagcac aaggttttta gttggagtac ctgaatatcg gaggtgggtt      420
309 gggaaatgat taccatcata ccgatgcagt cttacctaca cctatggatc tcatcaacac      480
310 tgtgcgagaa ttagttctct ctcaagatct cactcttatt attgaacccg gaagatcctt      540
311 gattgctaata acttgctgct tcgtcaatag agtaactggg gttaaatacta atggtacaaa      600
312 gaatttcatt gttgttgatg gcagcatggc agaactcatc agacctagtc tgtatggagc      660
313 ataccagcat atcgaaactgg tctctccccc cactcctggg gctgaagcag cgaccttcga      720
314 tattgttgga ccagtttgtg agtctgcaga tttccttgga aaagataggg aacttccaac      780
315 acctgatgag ggagctggac tggttgttca tgatgcaggc gcctactgca tgagcatggc      840
316 ttccacctac aacctgaagt tgaggccacc ggaatactgg gtggaagcgg acggttcgat      900
317 cgttaagatc aggcattggag agaagcttga tgactacatg aagttctttg atggtcttcc      960

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 03/06/2002  
PATENT APPLICATION: US/09/931,457A      TIME: 10:18:33

Input Set : A:\BB1116 US CIP Seq. list correct.txt  
Output Set: N:\CRF3\03062002\I931457A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 211,320,377,391,392,410,420,428,432,434,438,442,451  
Seq#:13; Xaa Pos. 128  
Seq#:16; N Pos. 373,406,430,433,455,494,504,553,579,583,595,596,620,639,644  
Seq#:16; N Pos. 650,661,673  
Seq#:18; N Pos. 465,524,537  
Seq#:19; Xaa Pos. 44  
Seq#:23; N Pos. 433,600  
Seq#:24; Xaa Pos. 56,57  
Seq#:27; N Pos. 271,421,425  
Seq#:35; N Pos. 260,306,376,383,404,432,446,455,456,509,514,522,537  
Seq#:39; N Pos. 400,417,486,492,493,505,518,524,530,532,557,563,581,591,596  
Seq#:39; N Pos. 617  
Seq#:40; Xaa Pos. 77,99

## VERIFICATION SUMMARY

DATE: 03/06/2002

PATENT APPLICATION: US/09/931,457A

TIME: 10:18:33

Input Set : A:\BB1116 US CIP Seq. list correct.txt

Output Set: N:\CRF3\03062002\I931457A.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:229 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:180  
L:231 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:300  
L:232 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:360  
L:233 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:420  
L:585 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:112  
L:804 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:360  
L:805 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:420  
L:806 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:480  
L:807 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:540  
L:808 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:600  
L:809 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:660  
L:863 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:420  
L:864 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18 after pos.:480  
L:885 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:32  
L:1055 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:420  
L:1057 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:540  
L:1080 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24 after pos.:48  
L:1247 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:240  
L:1250 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27 after pos.:420  
L:1648 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:240  
L:1649 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:300  
L:1650 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:360  
L:1651 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:420  
L:1652 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35 after pos.:480  
L:1895 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:360  
L:1897 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:480  
L:1898 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:540  
L:1899 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39 after pos.:600  
L:1929 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:64  
L:1935 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40 after pos.:96